Supplemental File for "Empirical comparisons of 12 meta-analysis methods for synthesizing proportions of binary outcomes"

Abbreviations/acronyms

Two-step (log): two-step method with the log transformation

Two-step (logit): two-step method with the logit transformation

Two-step (arcsine): two-step method with the arcsine-square-root transformation

Two-step (DAS-H): two-step method with the Freeman–Tukey double-arcsine transformation, using the harmonic mean of study-specific sample sizes as the overall sample size

Two-step (DAS-G): two-step method with the Freeman–Tukey double-arcsine transformation, using the geometric mean of study-specific sample sizes as the overall sample size

Two-step (DAS-A): two-step method with the Freeman–Tukey double-arcsine transformation, using the arithmetic mean of study-specific sample sizes as the overall sample size

Two-step (DAS-IV): two-step method with the Freeman–Tukey double-arcsine transformation, using the inverse of the variance of the synthesized result as the overall sample size

GLMM (log): generalized linear mixed model with the log link

GLMM (logit): generalized linear mixed model with the logit link

GLMM (probit): generalized linear mixed model with the probit link

GLMM (cauchit): generalized linear mixed model with the cauchit link

GLMM (cloglog): generalized linear mixed model with the complementary log-log link

AIC: Akaike information criterion

A. Statistical methods for meta-analysis of proportions

Suppose a meta-analysis of proportions contains N studies, and study i reports e_i events from n_i samples ($i = 1, \ldots, N$). The estimated event proportion in study i is $\hat{p}_i = e_i/n_i$. The two-step methods first apply a transformation to the estimated proportion with each study. The log transformation yields $g(\hat{p}_i) = \log \hat{p}_i$ with variance estimate $v_i = 1/e_i - 1/n_i$. The logit transformation yields $g(\hat{p}_i) = \log[\hat{p}_i/(1-\hat{p}_i)]$ with variance estimate $v_i = 1/e_i + 1/(n_i - e_i)$. The arcsine transformation yields $g(\hat{p}_i) = \arcsin\sqrt{\hat{p}_i}$ with variance estimate $v_i = 1/(4n_i)$. The Freeman–Tukey double-arcsine transformation yields $g(\hat{p}_i) = \arcsin\sqrt{e_i/(n_i+1)} + \arcsin\sqrt{(e_i+1)/(n_i+1)}$ with variance estimate $v_i = 1/(n_i + 0.5)$ (Freeman and Tukey, 1950). The log and logit transformations cannot directly handle zero event counts, and a continuity correction (e.g., 0.5) is needed for studies with zero event counts. The arcsine and Freeman–Tukey double-arcsine transformations do not need the continuity correction for zero event counts.

Consequently, conventional meta-analysis methods are applied to the study-specific transformed proportions $y_i = g(\hat{p}_i)$ and their variance v_i. Under the random-effects setting, the synthesized estimate is $\hat{\mu} = \sum_{i=1}^{N} w_i y_i / \sum_{i=1}^{N} w_i$ with variance $(\sum_{i=1}^{N} w_i)^{-1}$, where $w_i = 1/(v_i + \hat{\tau}^2)$ and $\hat{\tau}^2$ is an estimate of the between-study variance τ^2. Finally, $\hat{\mu}$ and its confidence interval are back-transformed to the original proportion scale (ranging from 0% to 100%).

The back-transformations of the log, logit, and arcsine transformations are straightforward; they are $g^{-1}(\hat{\mu}) = \exp(\hat{\mu})$, $g^{-1}(\hat{\mu}) = \exp(\hat{\mu})/[1 + \exp(\hat{\mu})]$, and $g^{-1}(\hat{\mu}) = \sin^2 \hat{\mu}$, accordingly. The back-transformation of the Freeman–Tukey double-arcsine transformation is complicated (Miller, 1978):

$$g^{-1}(\hat{\mu}; n) = 0.5 \left\{ 1 - \mathrm{sgn}(\cos \hat{\mu}) \sqrt{1 - \left(\sin \hat{\mu} + \frac{\sin \hat{\mu} - \sin^{-1} \hat{\mu}}{n} \right)^2} \right\},$$

where $\mathrm{sgn}(\cdot)$ denotes the sign function. This back-transformation depends on a specific sample size n. The sample size is explicitly defined in each individual study, but its definition is not clear for the synthesized result from the meta-analysis. Arguably, to represent a summary of the sample sizes of all studies in the meta-analysis, n may be the harmonic, geometric, or arithmetic mean of the study-specific sample sizes n_i. Of note, the harmonic mean $\leq$ the geometric mean $\leq$ the arithmetic mean.

Alternatively, the generalized linear mixed model (GLMM) is a one-step method for meta-analysis of proportions. It does not need data transformations at the individual study level. Instead, the GLMM directly models event counts with binomial likelihoods and uses a specific link function to transform latent true proportions to a linear scale. The GLMM can be specified as follows:

$$\begin{aligned} e_i &\sim bin(n_i, p_i); \\ g(p_i) &= \mu + \theta_i; \\ \theta_i &\sim N(0, \tau^2), \end{aligned}$$

where p_i is the true proportion of study i, μ is the overall proportion on the transformed scale via the link function $g(\cdot)$.

The various methods can be implemented with several R packages, including the functions `rma.uni()` and `rma.glmm()` in the R package "metafor," the function `metaprop()` in the R package "meta," and the functions `maprop.twostep()` and `maprop.glmm()` in the R package "altmeta." Users may refer to the manuals of these packages for instructions and examples.

References

Freeman, M. F. and Tukey, J. W. Transformations related to the angular and the square root. *The Annals of Mathematical Statistics*, 21(4):607–611, 1950.

Miller, J. J. The inverse of the freeman–tukey double arcsine transformation. *The American Statistician*, 32(4):138, 1978.

B. Additional results

Table S1: Basic characteristics of datasets of proportions from Cochrane reviews.

Characteristic	Median	Interquartile range (IQR)	Minimum	Maximum
Number of studies	5	3 to 7	3	167
Total event count	56	18 to 163	0	278,007
Total sample size	438	200 to 1,049	3	5,122,913
Crude event rate	12.6%	4.4% to 29.8%	0%	100%

Table S2: Number of datasets in categories based on the number of studies within a meta-analysis.

No. of studies	No. of datasets
3	13,059
4	8,173
5	5,428
6	3,671
7	2,616
8	2,018
9	1,553
10	1,208
11	922
12	719
13	583
14	502
15	413
16	329
17	255
18	251
19	195
20	149
$>$20	1,600

Table S3: Number of datasets in categories based on the total event count within a meta-analysis.

Total event count	No. of datasets
0	1,063
1	779
2	771
3	715
4	694
5	594
6	621
7	604
8	540
9	541
10	522
11	542
12	495
13	547
14	523
15	411
16	488
17	416
18	459
19	430
20	403
$>$20	31,486

Table S4: Number of datasets in categories based on the total sample size within a meta-analysis.

Total sample size	No. of datasets
1–10	33
11–20	45
21–30	82
31–40	226
41–50	356
51–60	484
61–70	586
71–80	586
81–90	685
91–100	698
101–110	742
111–120	644
121–130	789
131–140	811
141–150	747
151–160	669
161–170	706
171–180	629
181–190	740
191–200	706
$>$200	32,680

Table S5: Number of datasets in categories based on the crude event rate of a meta-analysis.

Crude event rate (%)	No. of datasets
0–0.1	1,211
0.1–0.5	1,040
0.5–1	1,272
1–5	8,379
5–10	7,110
10–20	8,719
20–30	5,084
30–40	3,276
40–50	2,515
50–60	1,752
60–70	1,361
70–80	963
80–90	610
90–95	178
95–99	144
99–99.5	13
99.5–99.9	3
99.9–100	14

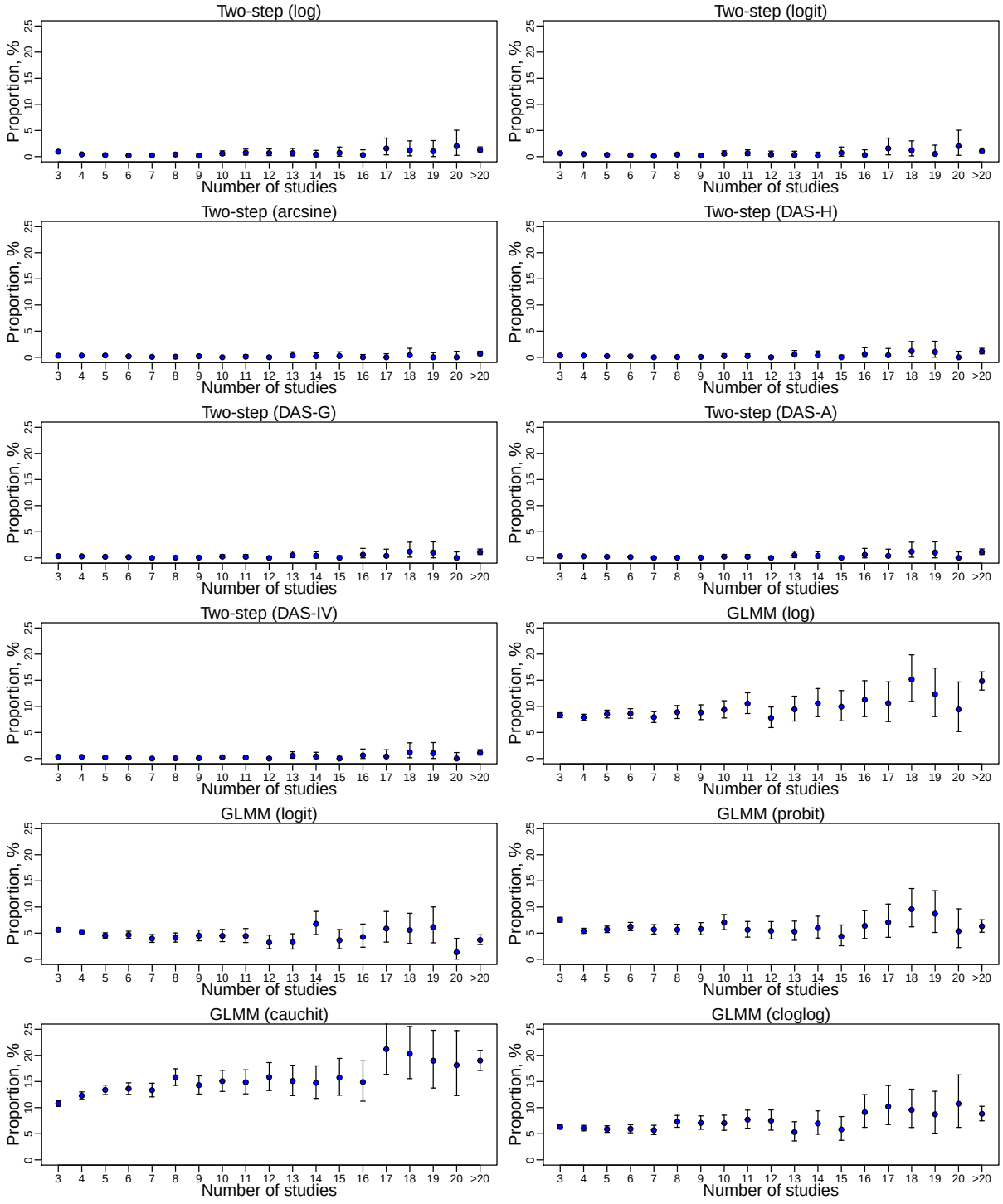

Figure S1: Proportions (with 95% confidence intervals) of Cochrane datasets that led to computational issues when using various meta-analysis methods, categorized by the number of studies within a meta-analysis.

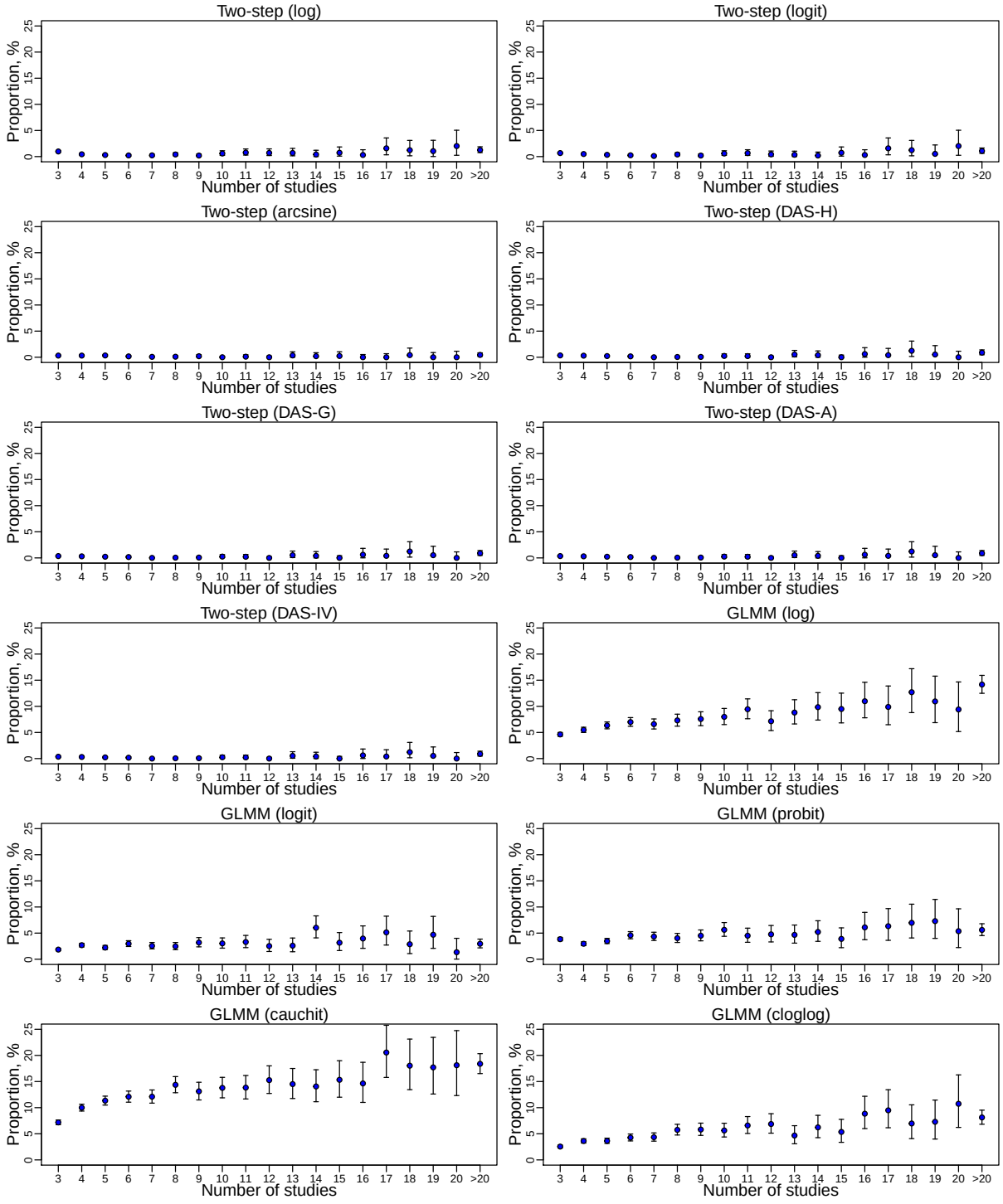

Figure S2: Proportions (with 95% confidence intervals) of Cochrane datasets that led to computational issues when using various meta-analysis methods, categorized by the number of studies within a meta-analysis, among the datasets with at least 1 total event count.

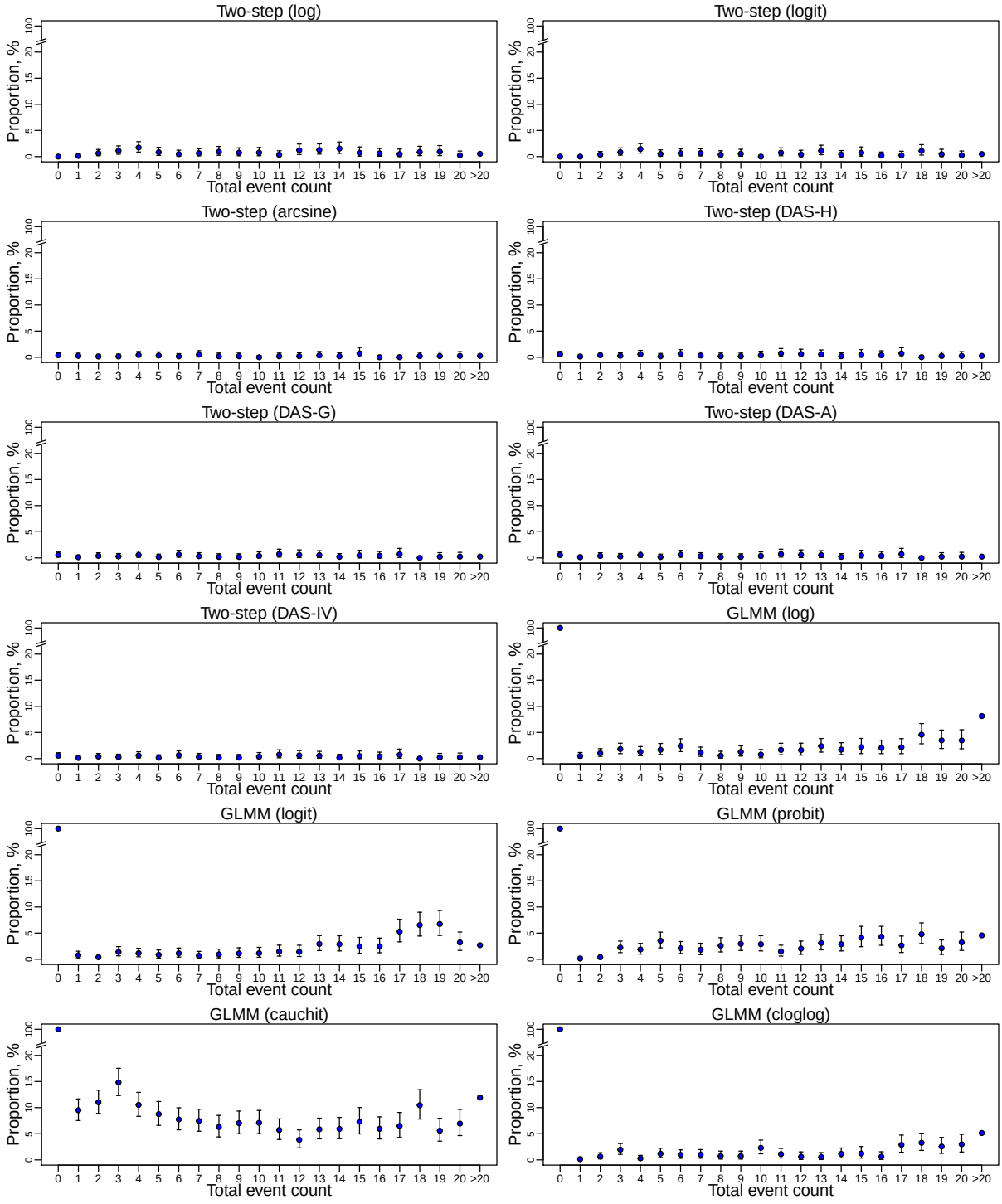

Figure S3: Proportions (with 95% confidence intervals) of Cochrane datasets that led to computational issues when using various meta-analysis methods, categorized by the total event count within a meta-analysis.

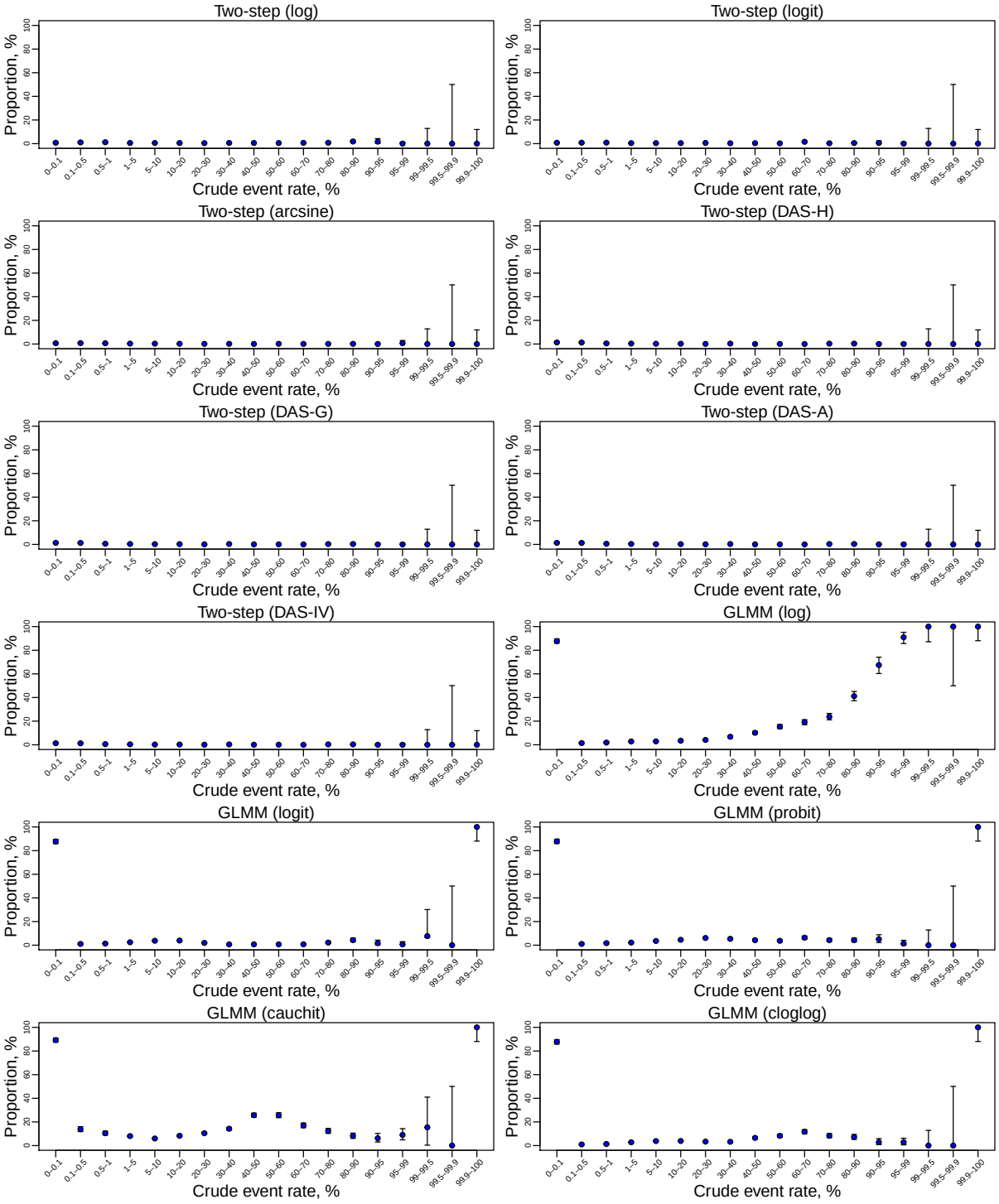

Figure S4: Proportions (with 95% confidence intervals) of Cochrane datasets that led to computational issues when using various meta-analysis methods, categorized by the crude event rate of a meta-analysis.

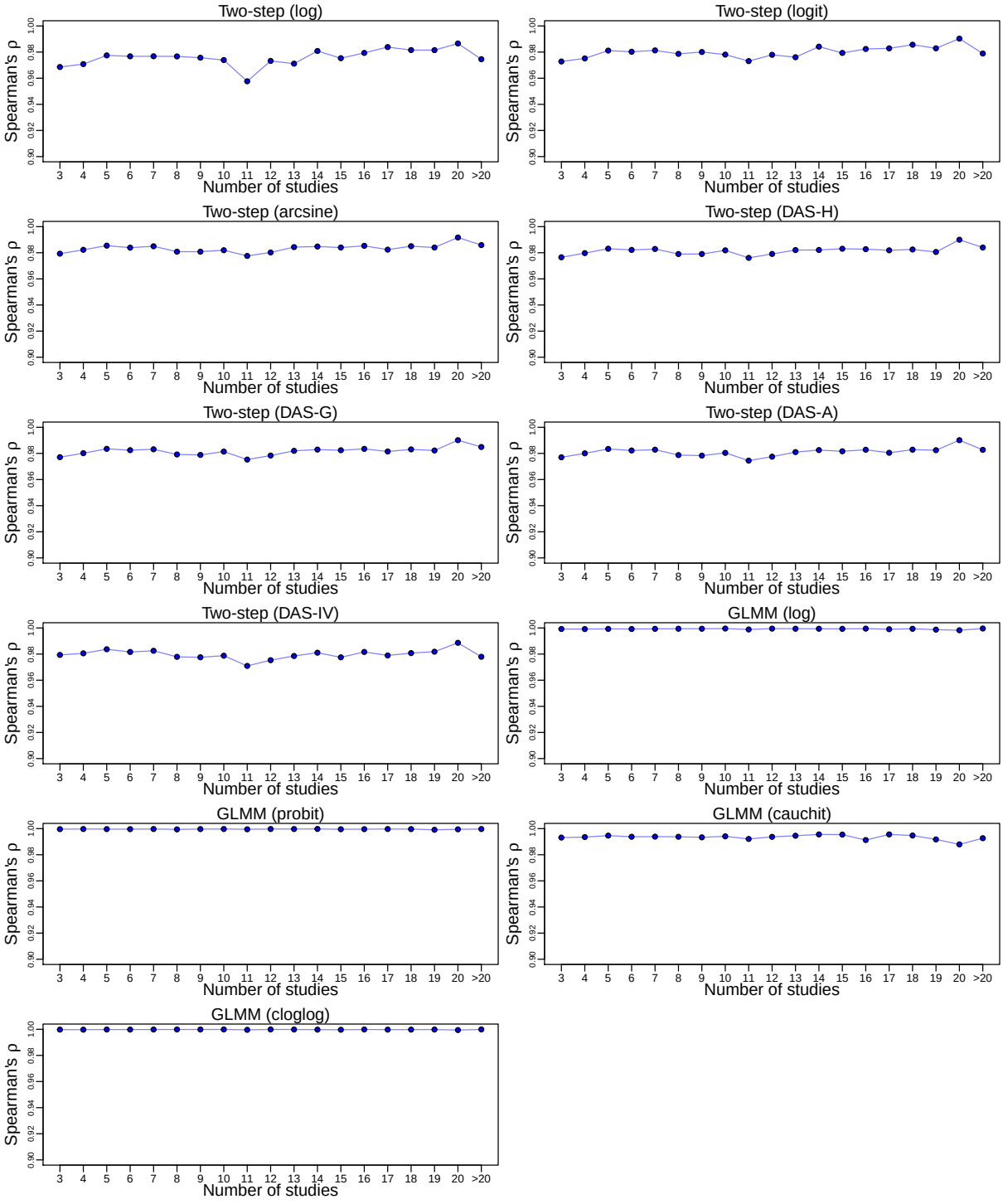

Figure S5: Spearman's rank correlation coefficients ρ between overall proportion estimates produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the number of studies within a meta-analysis.

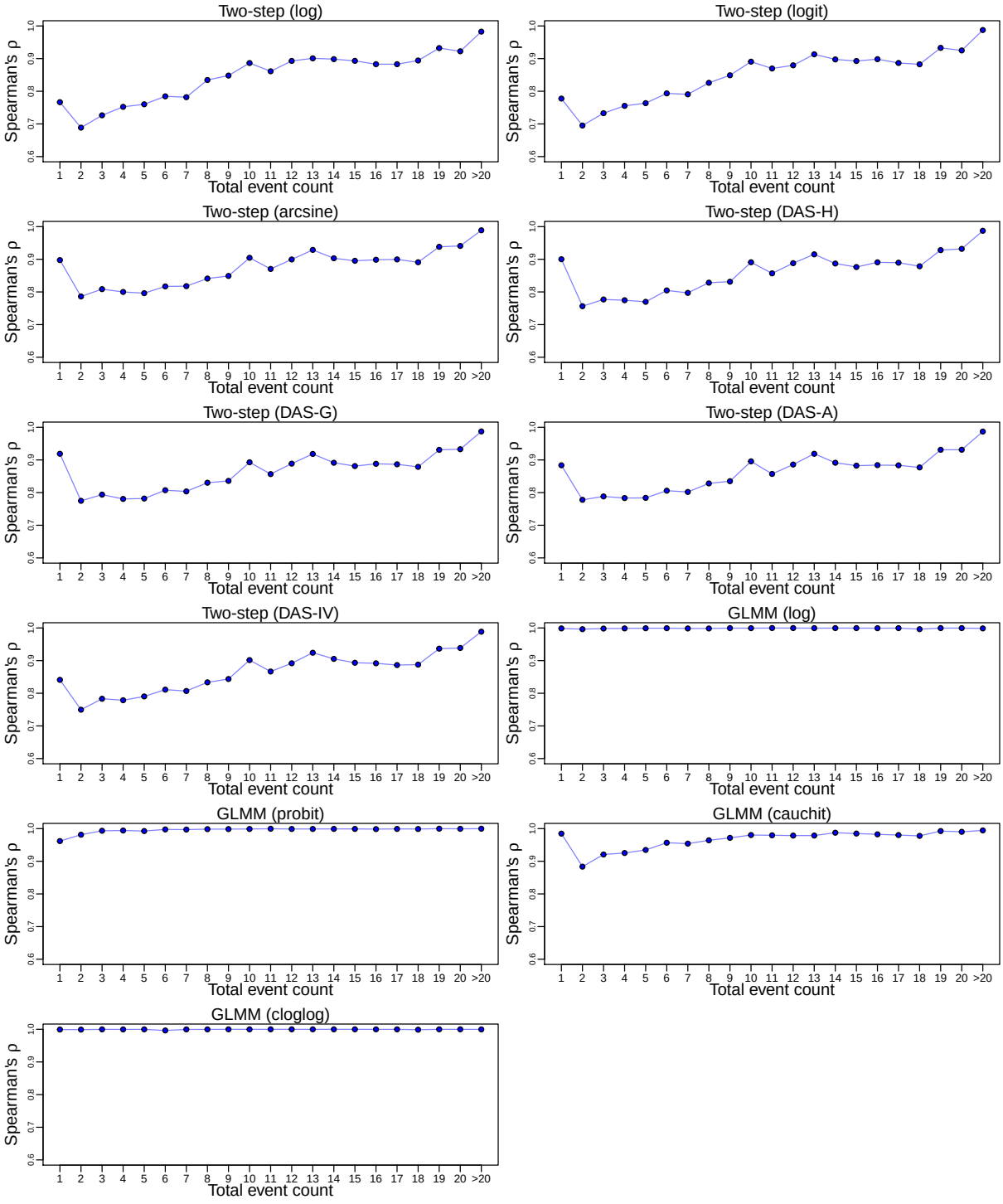

Figure S6: Spearman's rank correlation coefficients ρ between overall proportion estimates produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the total event count within a meta-analysis.

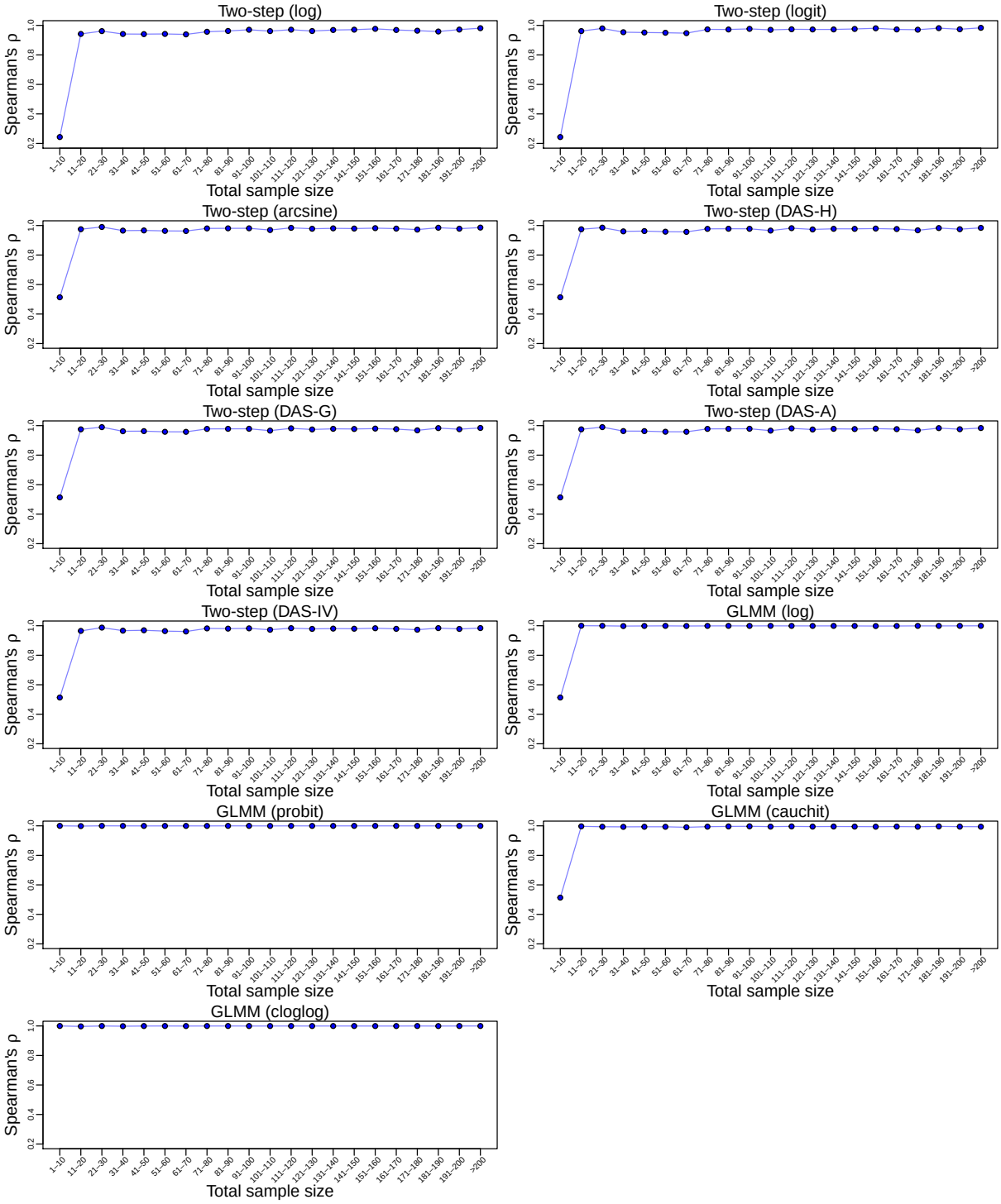

Figure S7: Spearman's rank correlation coefficients ρ between overall proportion estimates produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the total sample size within a meta-analysis.

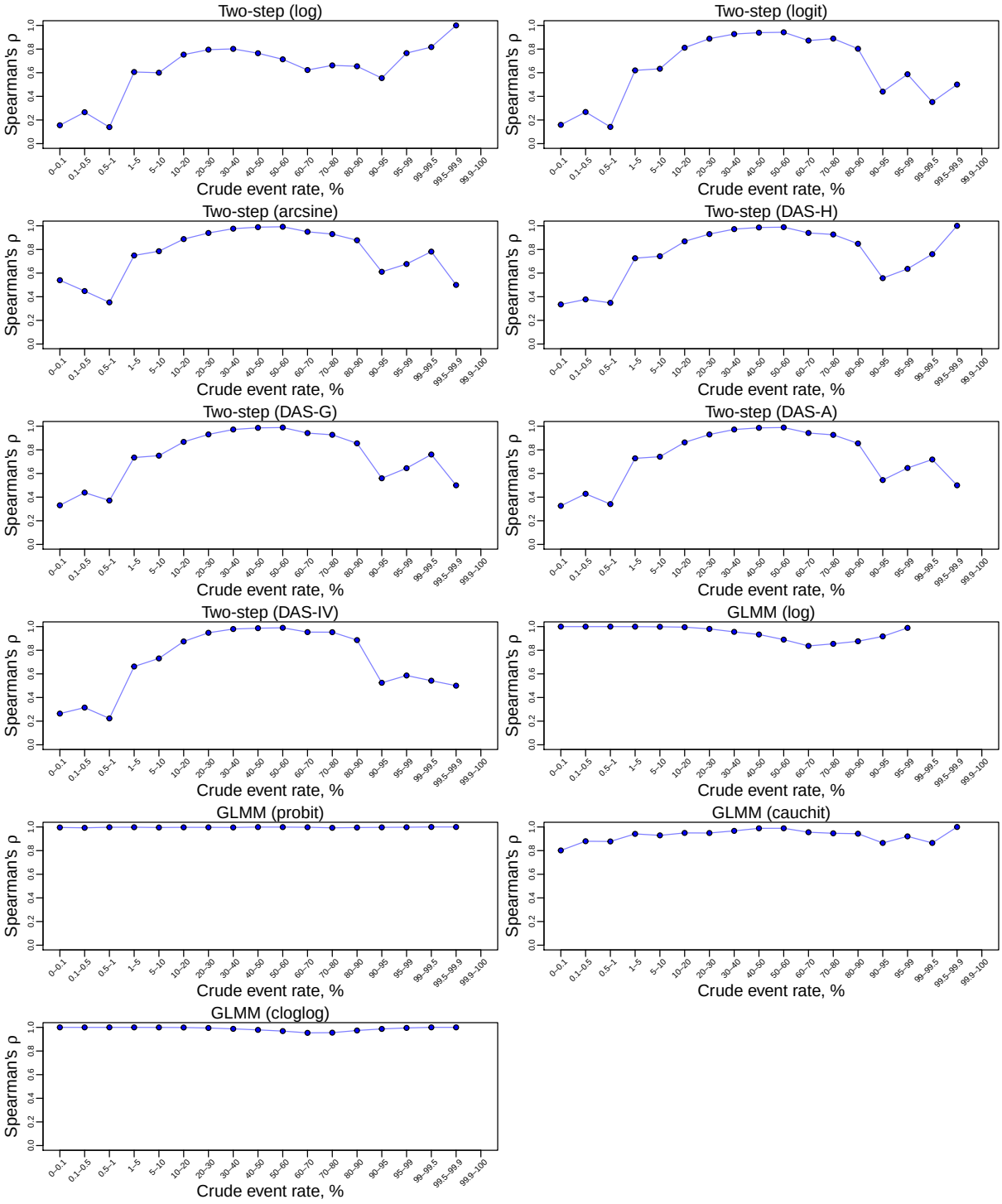

Figure S8: Spearman's rank correlation coefficients ρ between overall proportion estimates produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the crude event rate of a meta-analysis.

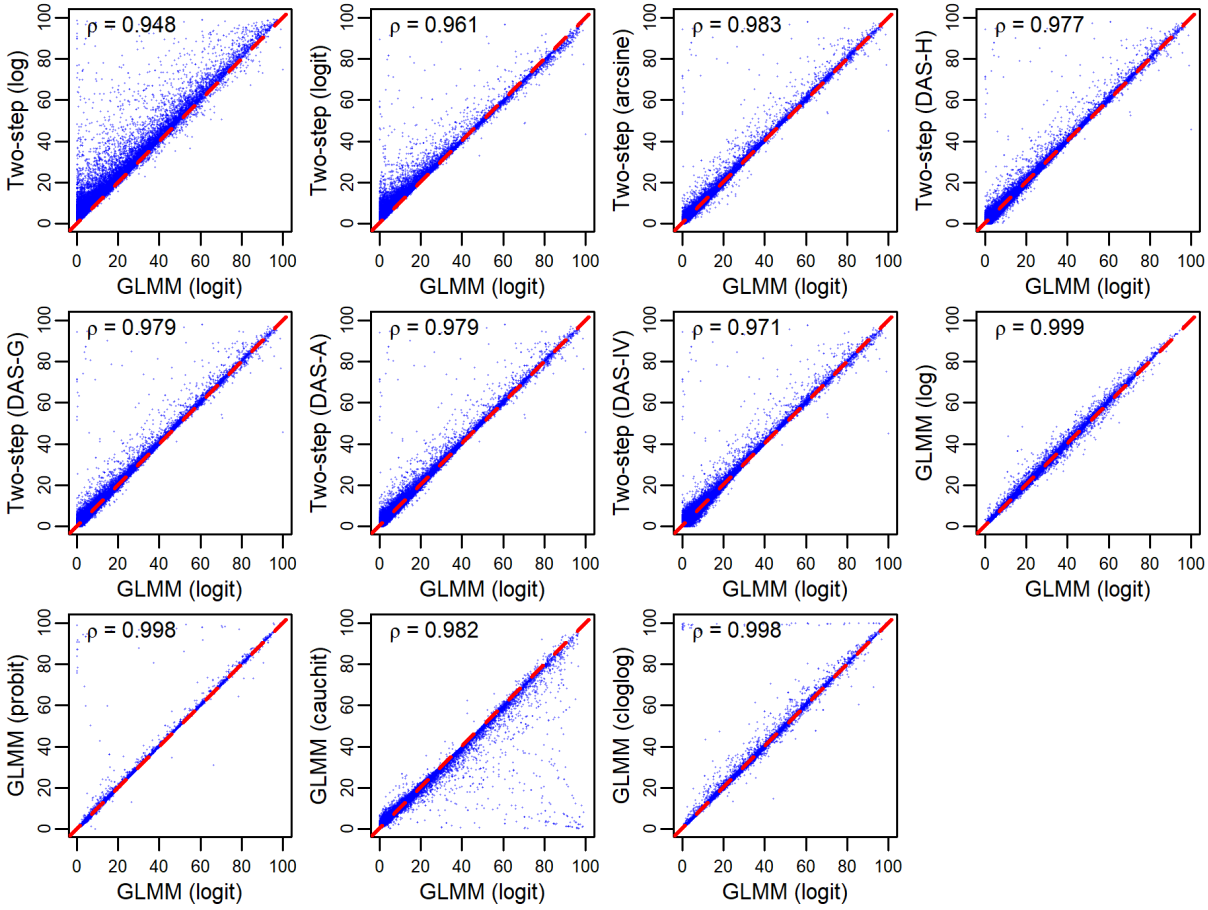

Figure S9: Scatter plots of 95% confidence interval lower bounds of overall proportion estimates (in percentage) produced by various meta-analysis methods against those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, the diagonal dashed line represents the identity line, and Spearman's rank correlation coefficient ρ between the corresponding two meta-analysis methods is displayed.

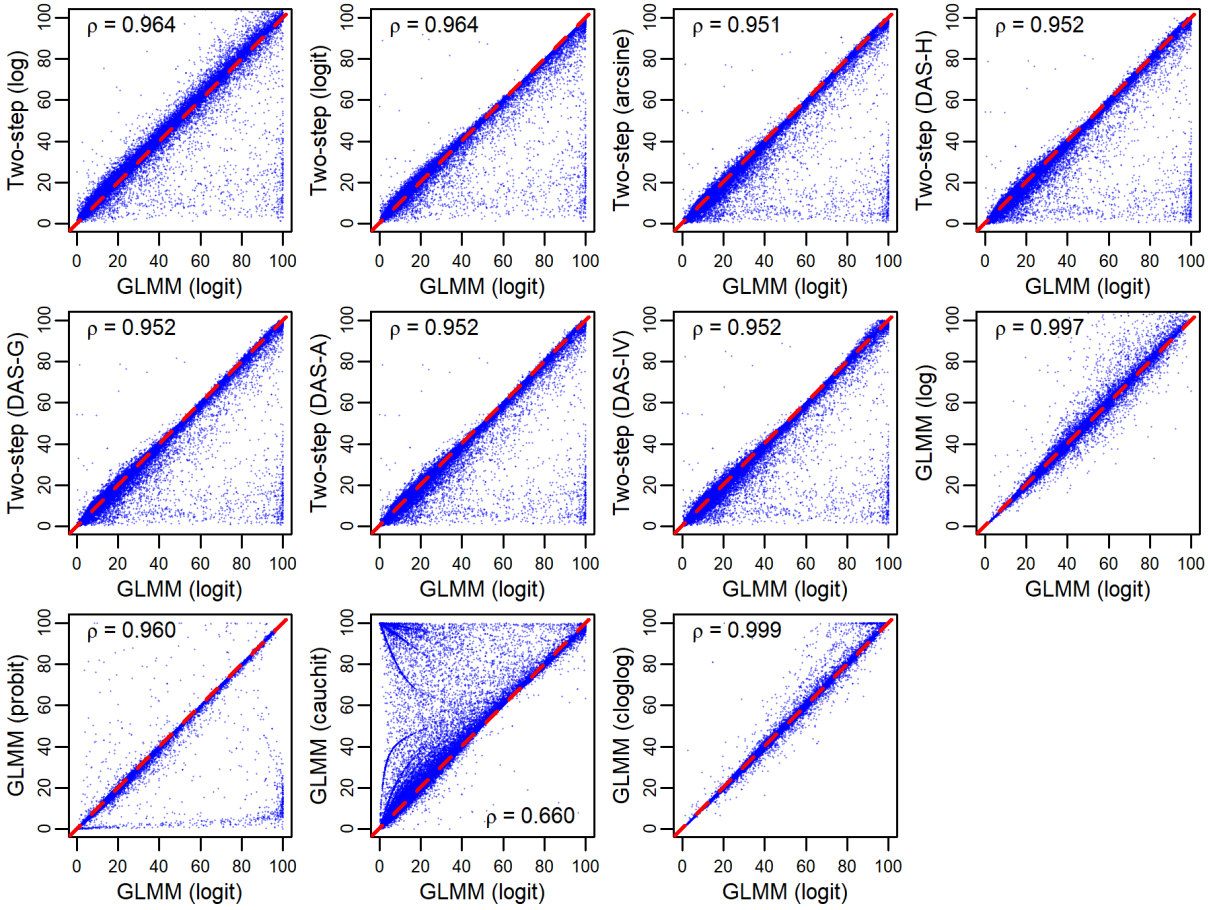

Figure S10: Scatter plots of 95% confidence interval upper bounds of overall proportion estimates (in percentage) produced by various meta-analysis methods against those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, the diagonal dashed line represents the identity line, and Spearman's rank correlation coefficient ρ between the corresponding two meta-analysis methods is displayed. Of note, 95% confidence interval upper bounds produced by the two-step method with the log transformation and those produced by the generalized linear mixed model with the log link may exceed 100%; they were not truncated in our analyses for calculating Spearman's ρ, and the corresponding two panels only present the incomplete results of upper bounds within 0%–100%.

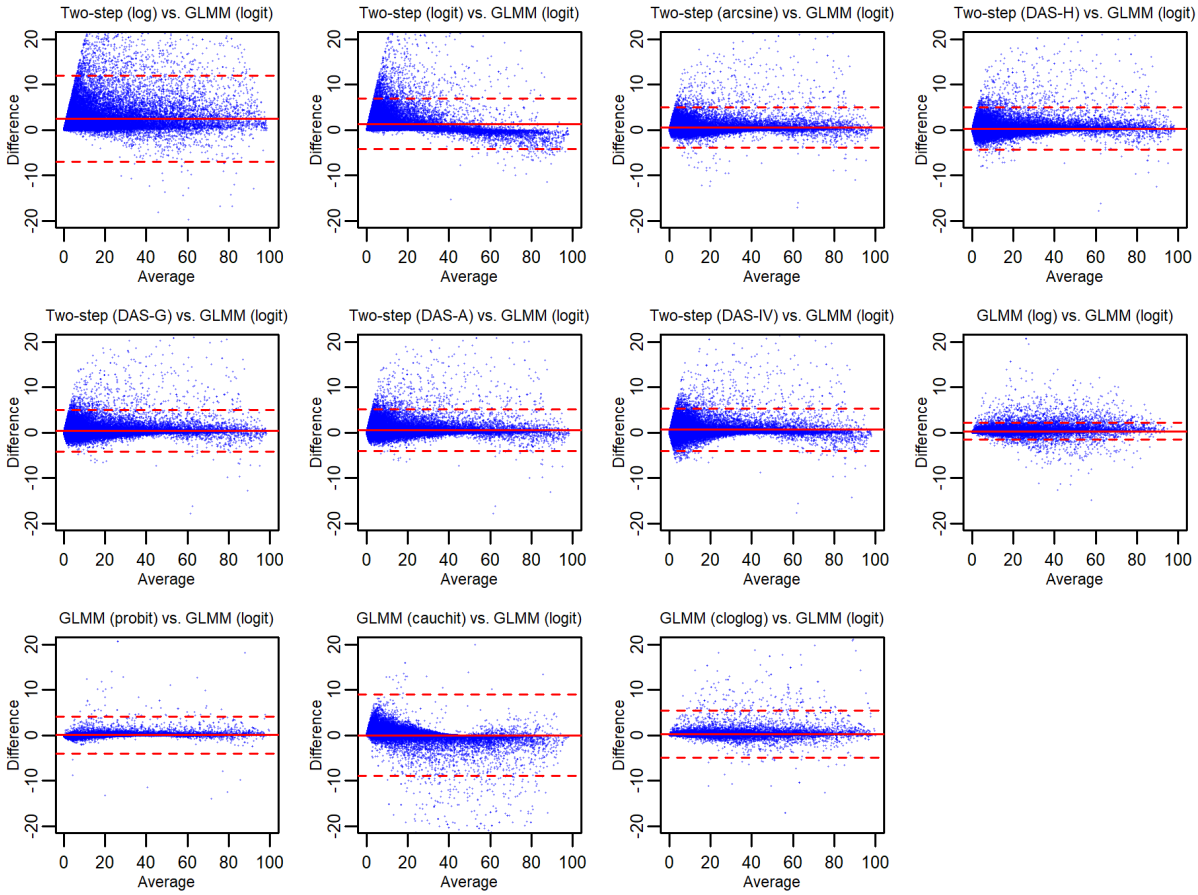

Figure S11: Bland–Altman plots of agreement between 95% confidence interval lower bounds of overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, the horizontal solid line represents the mean difference, and the horizontal dashed lines represent 95% limits of agreement.

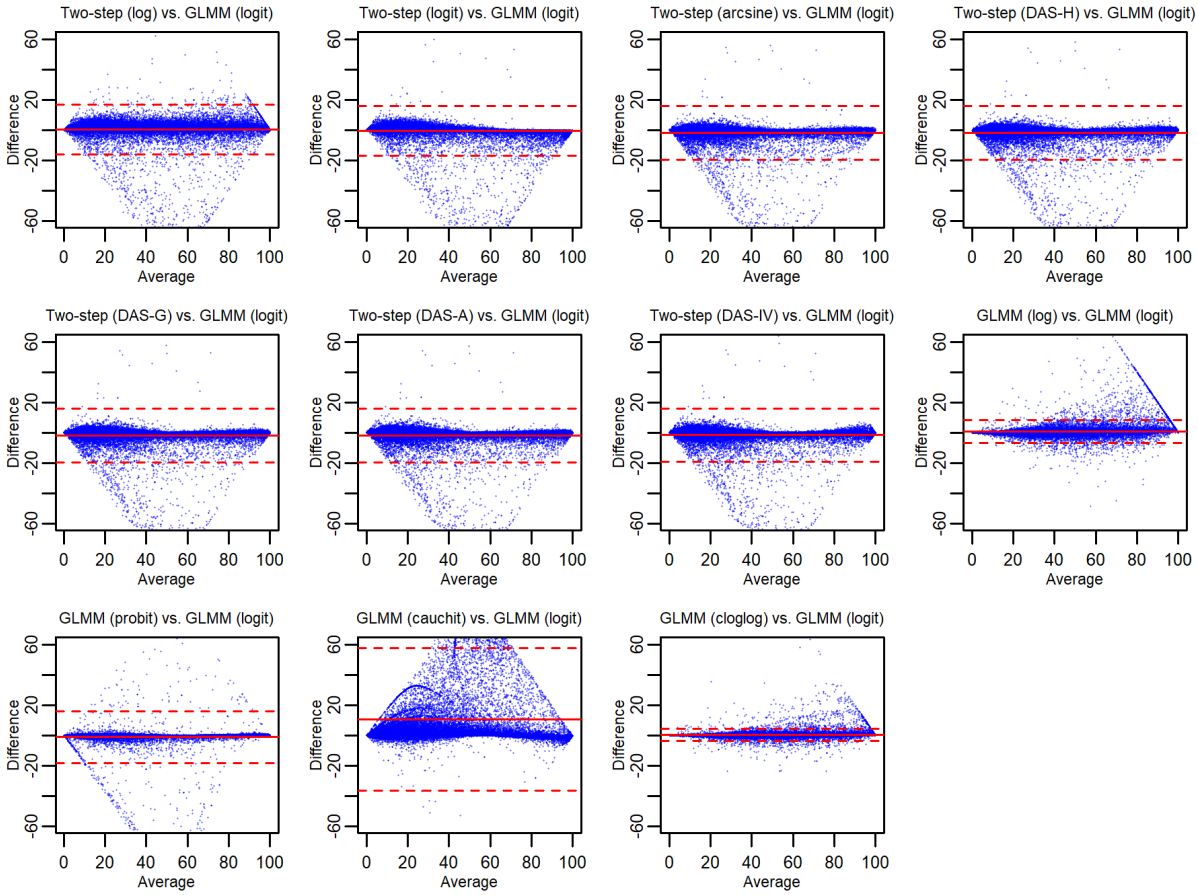

Figure S12: Bland–Altman plots of agreement between 95% confidence interval upper bounds of overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, the horizontal solid line represents the mean difference, and the horizontal dashed lines represent 95% limits of agreement.

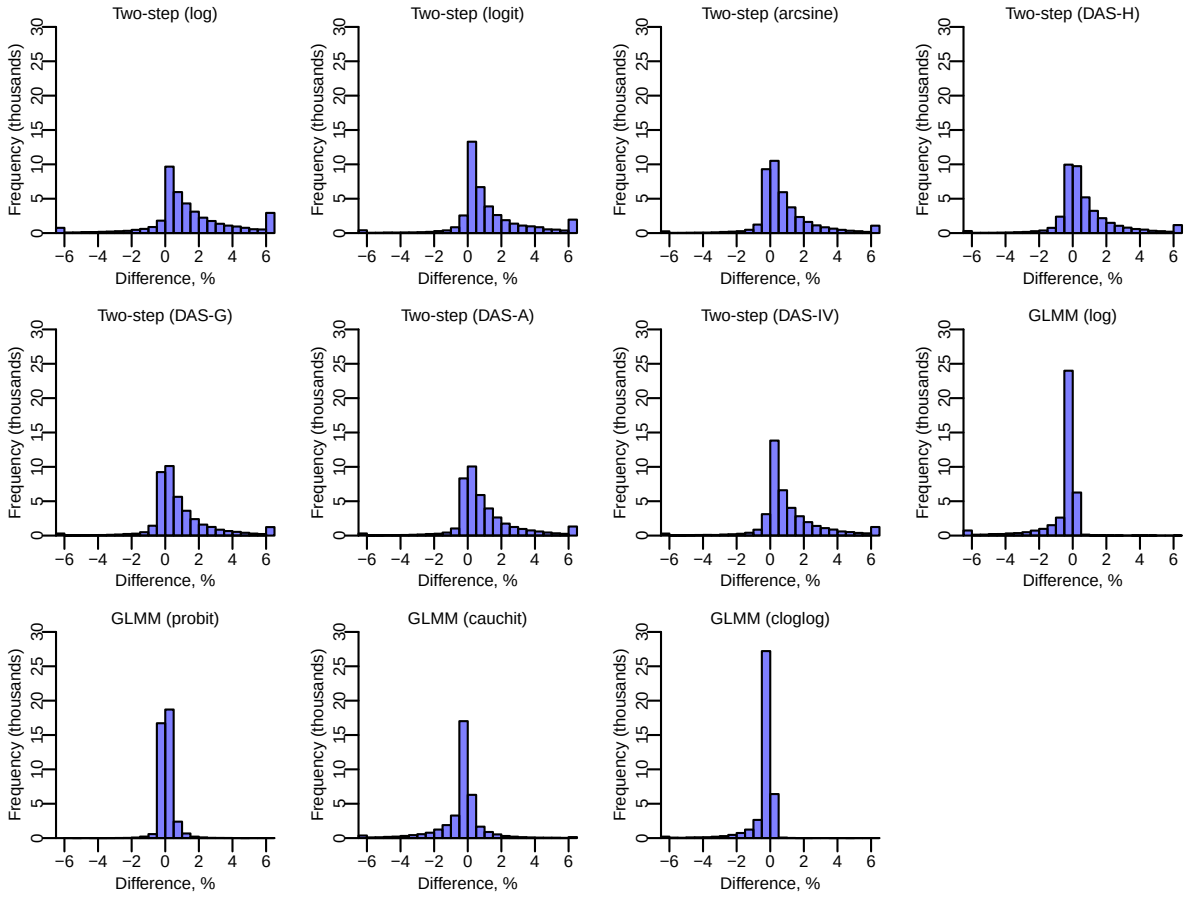

Figure S13: Histograms of absolute differences between overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, each bar's width is 0.5%; the bar to the left of -6% (if any) represents absolute differences truncated at -6%, and the bar to the right of 6% (if any) represents absolute differences truncated at 6%.

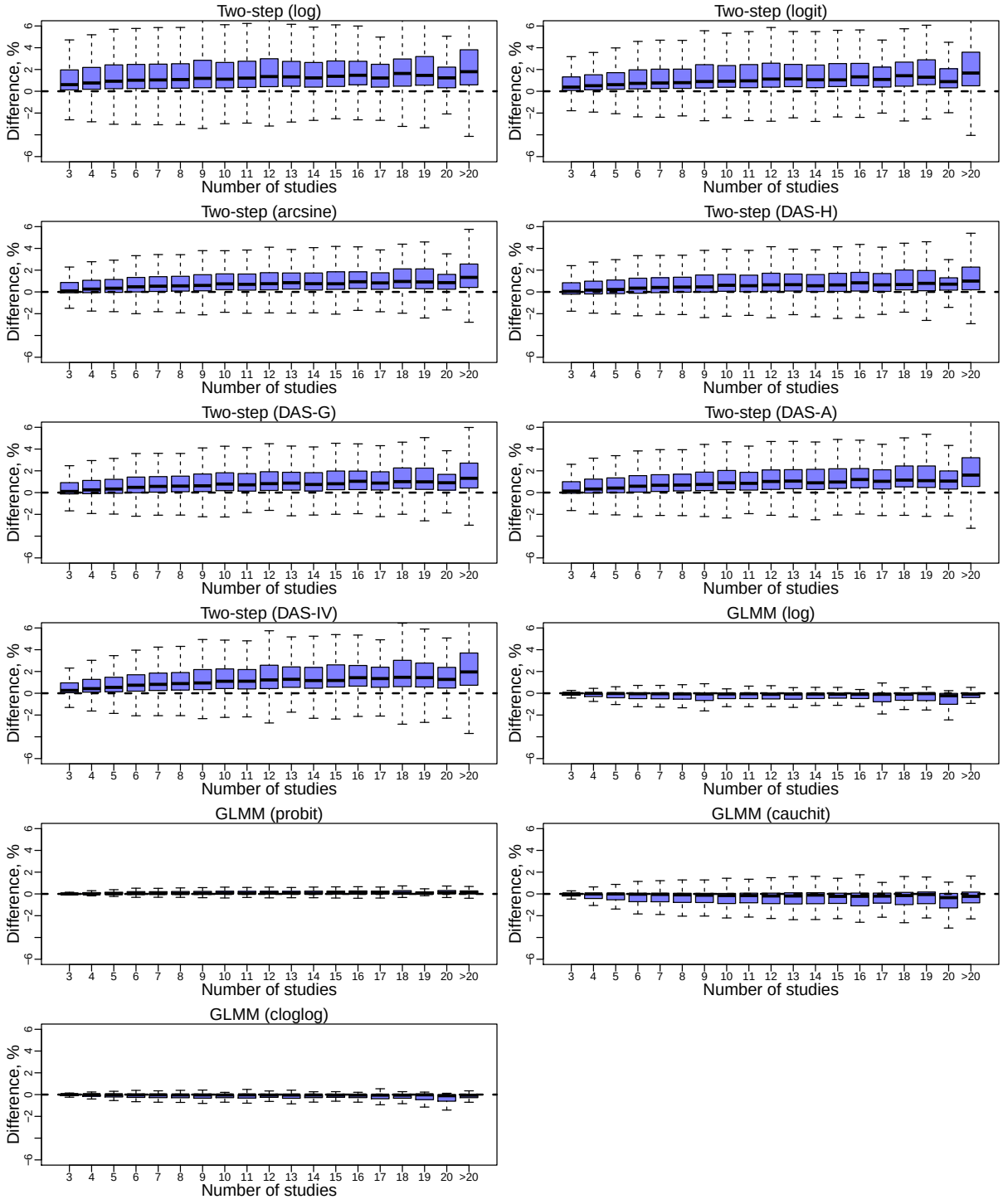

Figure S14: Box plots of absolute differences between overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the number of studies within a meta-analysis. In each panel, the horizontal dashed line represents no absolute difference. Because many datasets led to extreme values of absolute differences, the box plots do not present outliers, and the vertical axis only presents the range from -6% to 6%.

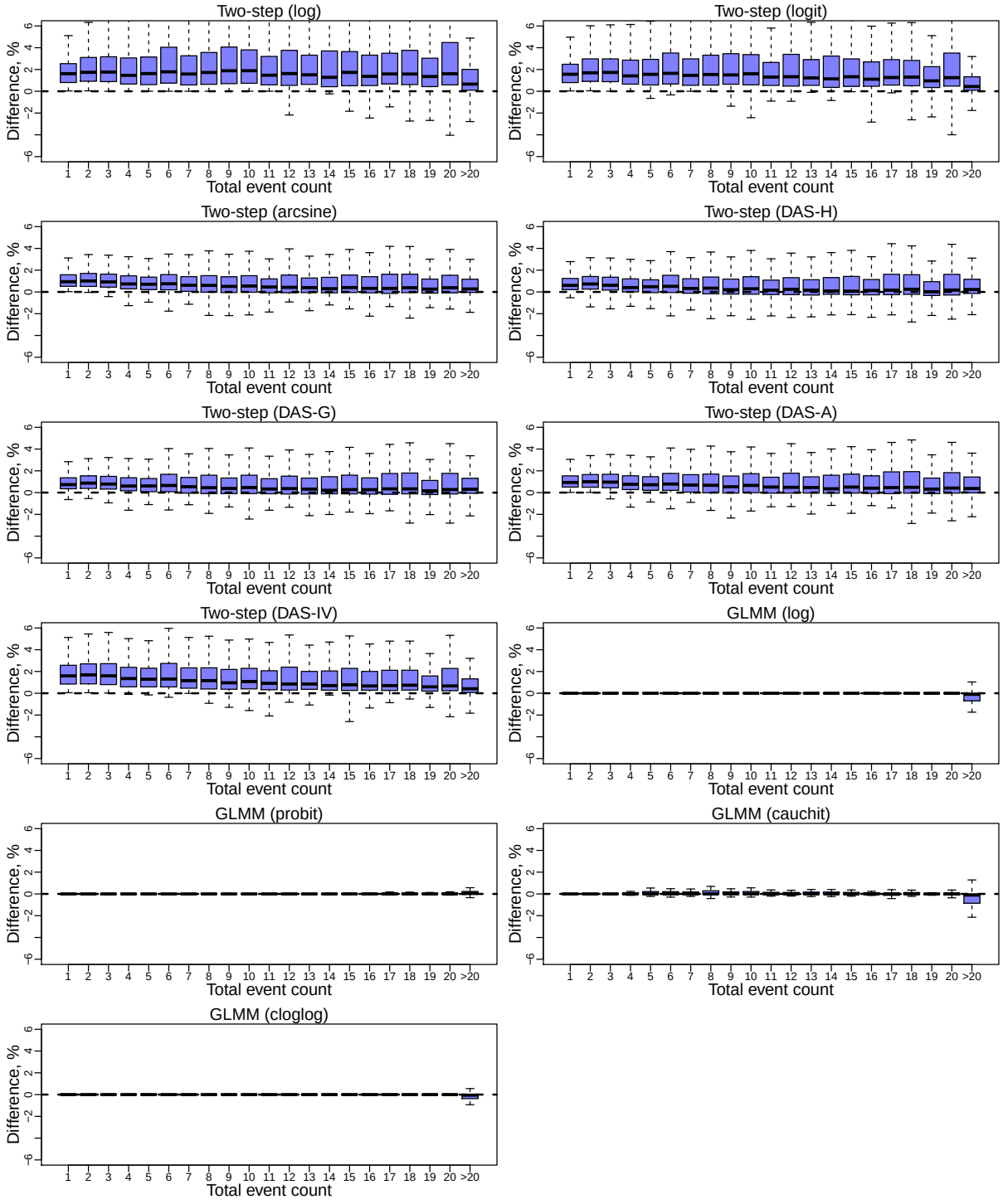

Figure S15: Box plots of absolute differences between overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the total event count within a meta-analysis. In each panel, the horizontal dashed line represents no absolute difference. Because many datasets led to extreme values of absolute differences, the box plots do not present outliers, and the vertical axis only presents the range from -6% to 6%.

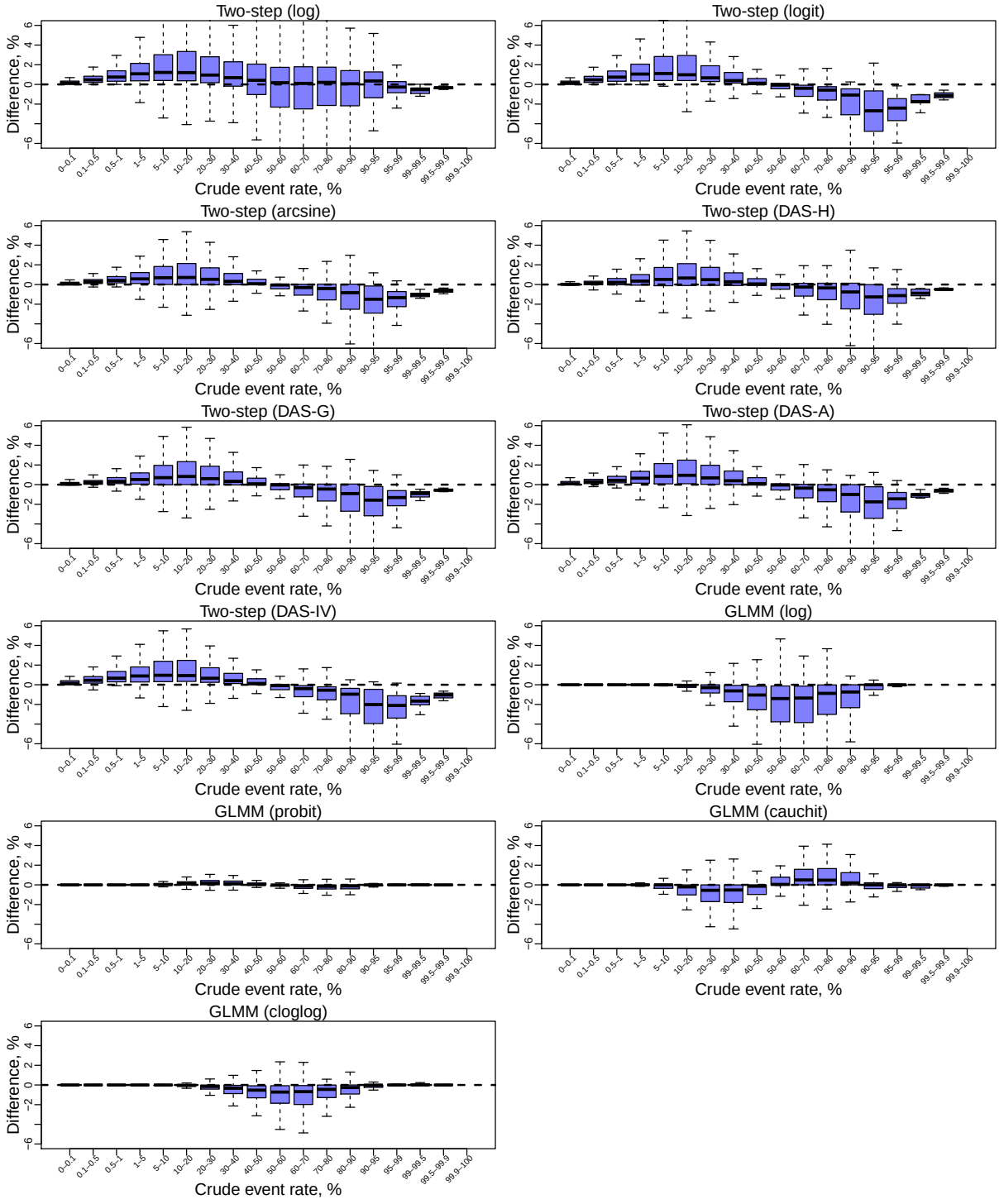

Figure S16: Box plots of absolute differences between overall proportion estimates (in percentage) produced by various meta-analysis methods and those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the crude event rate of a meta-analysis. In each panel, the horizontal dashed line represents no absolute difference. Because many datasets led to extreme values of absolute differences, the box plots do not present outliers, and the vertical axis only presents the range from -6% to 6%.

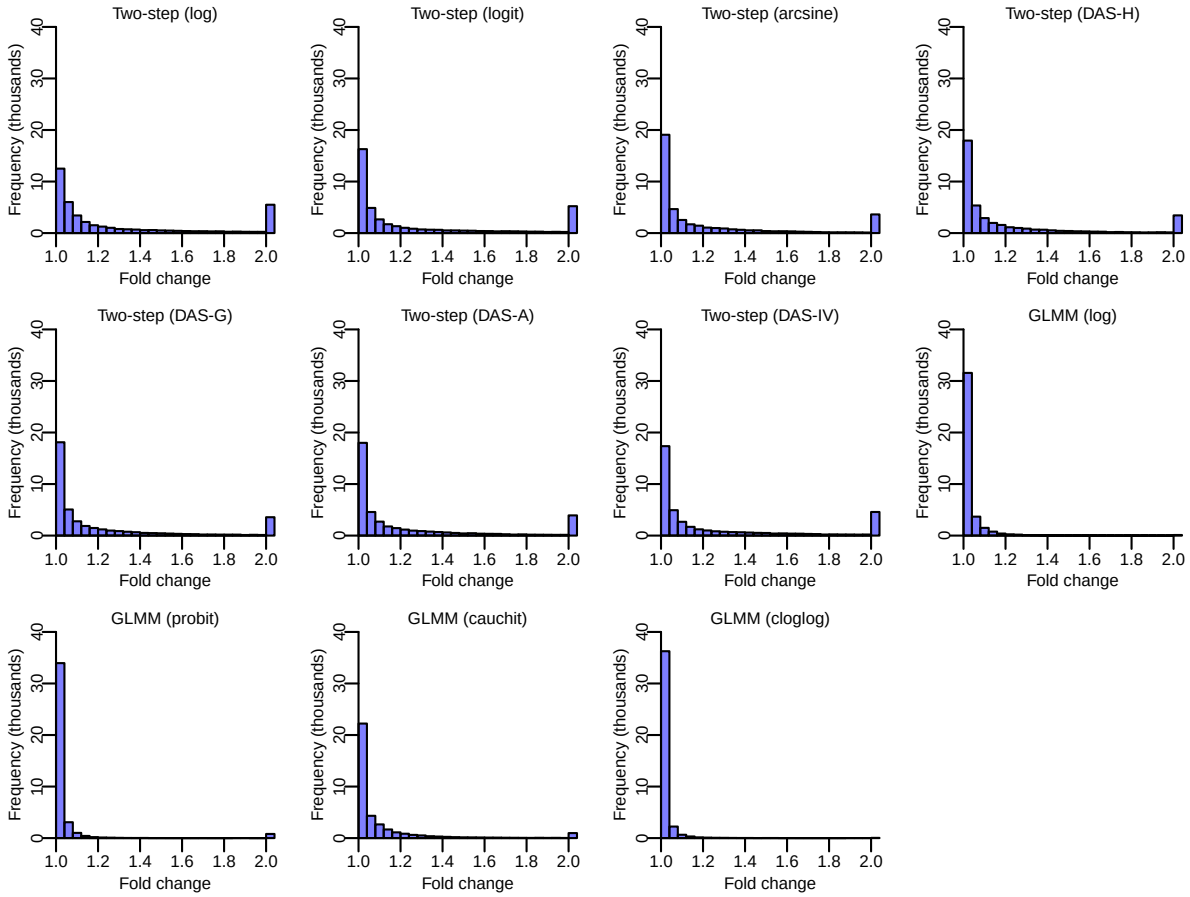

Figure S17: Histograms of fold changes of overall proportion estimates produced by various meta-analysis methods compared with those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, each bar's width is 0.04; the bar to the right of 2 (if any) represents fold changes truncated at 2.

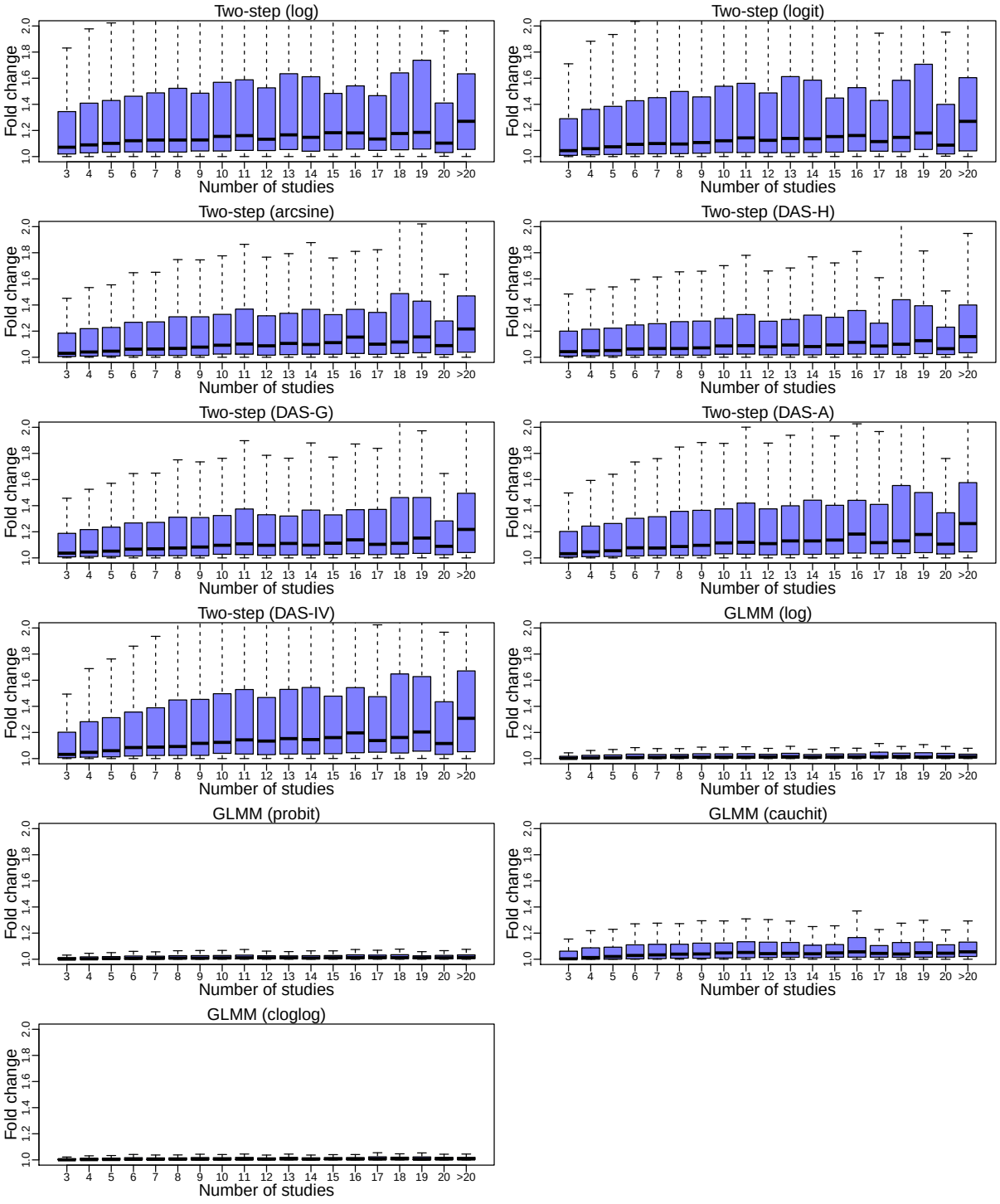

Figure S18: Box plots of fold changes of overall proportion estimates produced by various meta-analysis methods compared with those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the number of studies within a meta-analysis. Because many datasets led to extreme values of fold changes, the box plots do not present outliers, and the vertical axis only presents the range from 1 to 2.

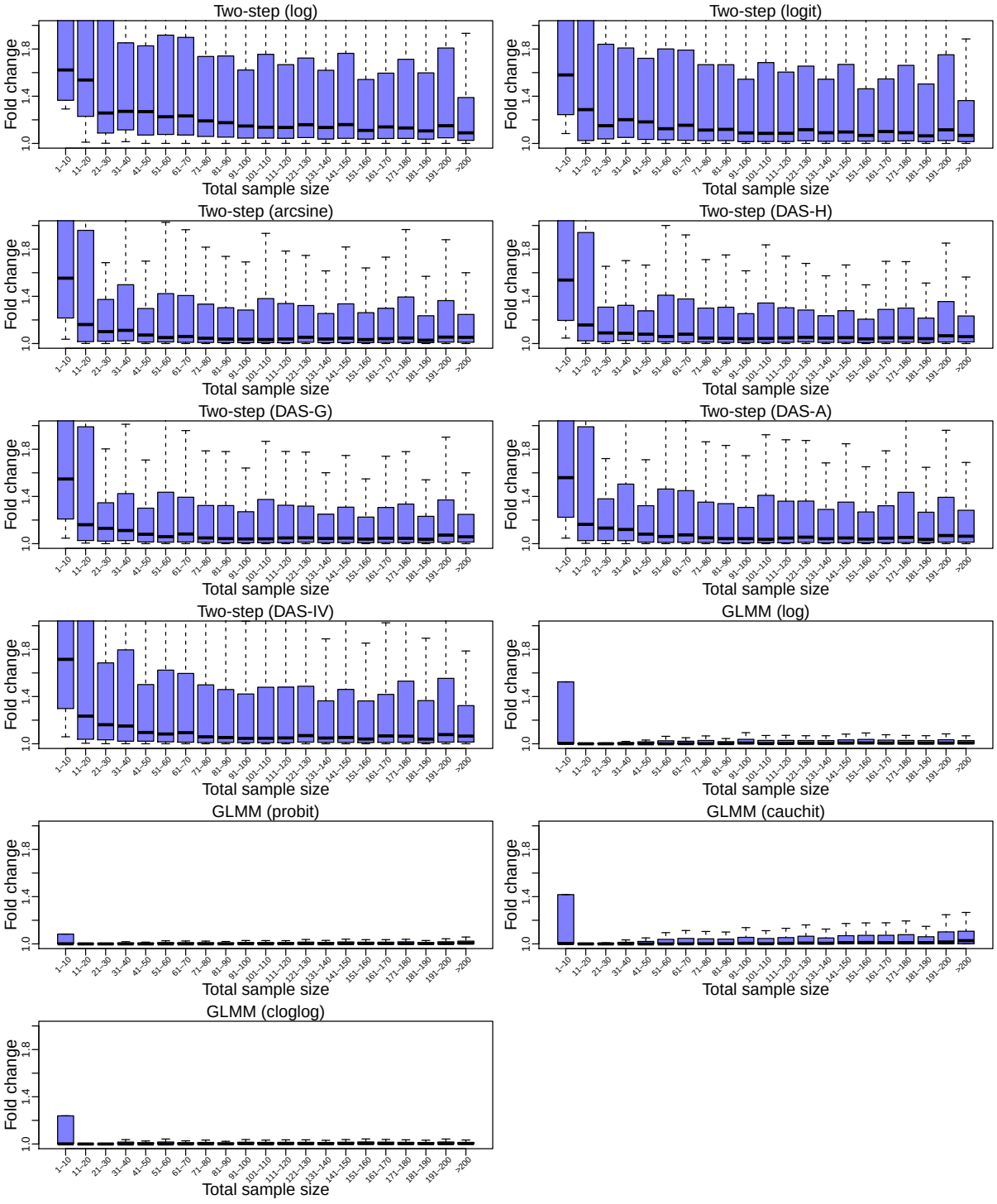

Figure S19: Box plots of fold changes of overall proportion estimates produced by various meta-analysis methods compared with those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the total sample size within a meta-analysis. Because many datasets led to extreme values of fold changes, the box plots do not present outliers, and the vertical axis only presents the range from 1 to 2.

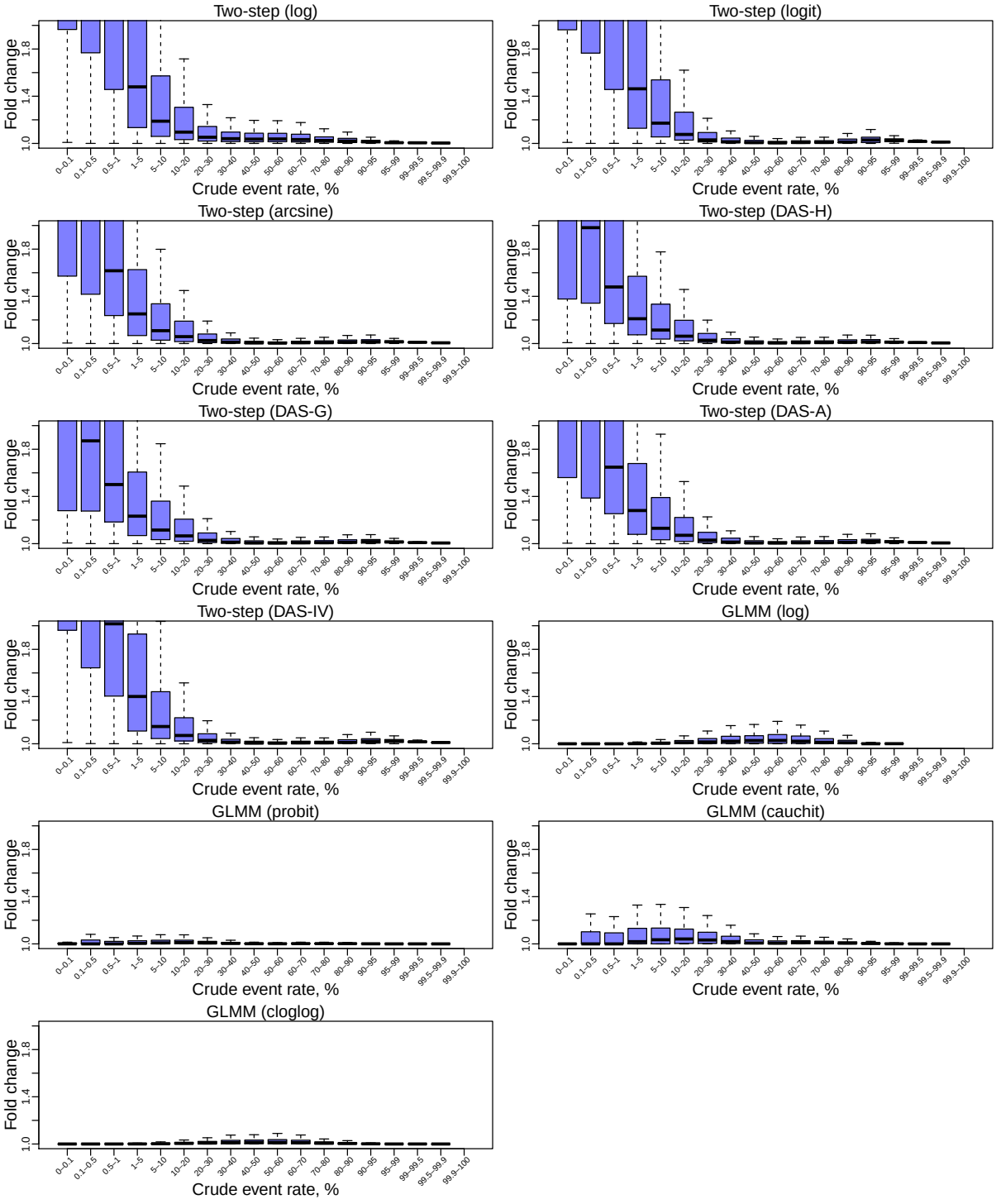

Figure S20: Box plots of fold changes of overall proportion estimates produced by various meta-analysis methods compared with those by the generalized linear mixed model with the logit link among Cochrane datasets, categorized by the crude event rate of a meta-analysis. Because many datasets led to extreme values of fold changes, the box plots do not present outliers, and the vertical axis only presents the range from 1 to 2.

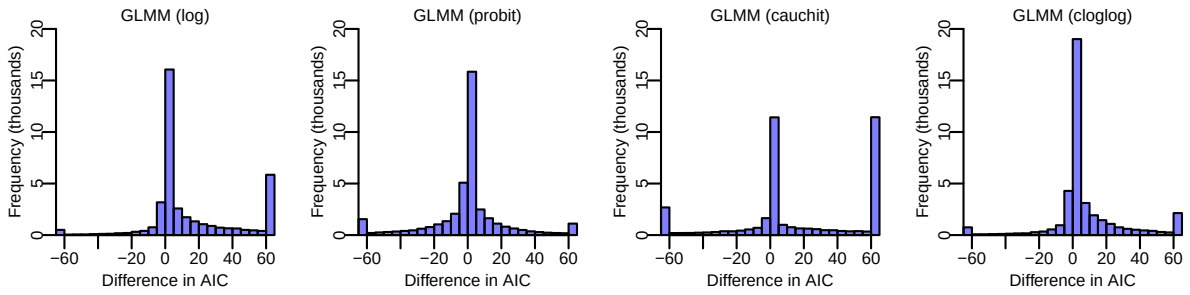

Figure S21: Histograms of differences in AIC values produced by generalized linear mixed models with various link functions compared with those by the generalized linear mixed model with the logit link among Cochrane datasets. In each panel, each bar's width is 5; the bar to the left of -60 (if any) represents differences in AIC truncated at -60, and the bar to the right of 60 (if any) represents differences in AIC truncated at 60.